СЕРТИФИКАЦИЯ ЭКСПЕРТОВ И ОПРЕДЕЛЕНИЕ ОТНОСИТЕЛЬНОЙ ЦЕНЫ ЗАДАЧИ В ЗАВИСИМОСТИ ОТ ЕЕ СЛОЖНОСТИ

В.И. Протасов, З.Е. Потапова, О.К. Осипчук
(Москва, МАИ)

CERTIFICATION OF EXPERTS AND DETERMINATION OF RELATIVE PRICE OF THE TASK

V. Protasov, Z. Potapova, O. Osiptchuk
(Moscow Aviation Institute)

Within the Rash's model the relative price of a task is determined depending on its complexity. For this purpose the concept of the ideal expert is entered. This approach is developed as an extension of the Evolutionary Solutions Coordination method. The method of simultaneous determination of complexity of test tasks and kompetentnost of experts that allows to carry out certification of experts is received. Computer models allow to calculate the relative cost of work of experts for a problem of the set complexity and the group of experts of identical readiness. Conditions for an optimum choice of quantity and competence of experts for the guaranteed solution of a task with known complexity are found.

I . Введение. В настоящее время наблюдается экспоненциальный рост числа научных публикаций, посвященных коллективному интеллекту. С использованием социального WEB-2 и социального компьютинга созданы и используются новые сетевые инструменты, такие, например, как краудсорсинг (англ. crowdsourcing, crowd — «толпа» и sourcing — «использование ресурсов») — применение данного метода подразумевает делегирование бизнес-задачи фирмы-организатора удаленному сетевому сообществу. Такой способ организации труда имеет существенные преимущества перед традиционным наймом сотрудников, поскольку позволяет быстро создавать глобальный продукт при использовании дешевой удаленной рабочей силы. При этом бизнес-риски разделяются с исполнителями, которые в большинстве случаев получают оплату своего труда после продажи продукта. Впервые термин «crowdsourcing» использовал в 2006 году журналист Джефф Хауи[1]. В России данную технологию активно развивает недавно созданная фирма Witology, осуществившая ряд значимых для российской экономики проектов [2].

Одной из основных проблем краудсорсинга является непредсказуемость его результатов, связанная с тем обстоятельством, что по теореме Кондорсе[3], положенной в основу этого метода, требуется, чтобы вероятность правильного заключения эксперта превышала 0.5 и для надежности метода требуется предварительное тестирование большого количества экспертов. Дело усугубляется тем, что в реальной практике слоты (составные части) проекта могут иметь различную сложность, и она должна учитываться при подготовке и прогнозировании выполнения проекта.

Открытыми на сегодняшний день также остаются вопросы определения стоимости интеллектуальной работы в зависимости от ее сложности, а также вопросы справедливой оплаты труда отдельных экспертов и групп экспертов при решении интеллектуальных задач.

II. Модель Раша. Связь между уровнем сложности тестовых вопросов и степенью подготовленности экспертов при определении вероятности правильного ответа была установлена в наиболее общей теории конструирования тестов, опирающейся на теорию измерения, – Item Response Theory (IRT)[4]. Для наших целей подходящей является однопараметрическая модель Раша[5-8], как наиболее простая модель, связывающая вероятность получения правильного ответа G испытуемого с уровнем его подготовленности (компетентности) θ и сложностью задания β :

$$G = \frac{1}{1+e^{\alpha(\beta-\theta)}}, \qquad (1)$$

где α – масштабный множитель. Данное выражение представляет собой логистическую кривую, уровни сложности задачи и подготовленности специалистов измеряются в специальных единицах – логитах.

На рис. 1 показаны кривые Раша при *α=1* для сложности задания β от –3 логит (самое легкое задание) до 3 логит (самое трудное задание).

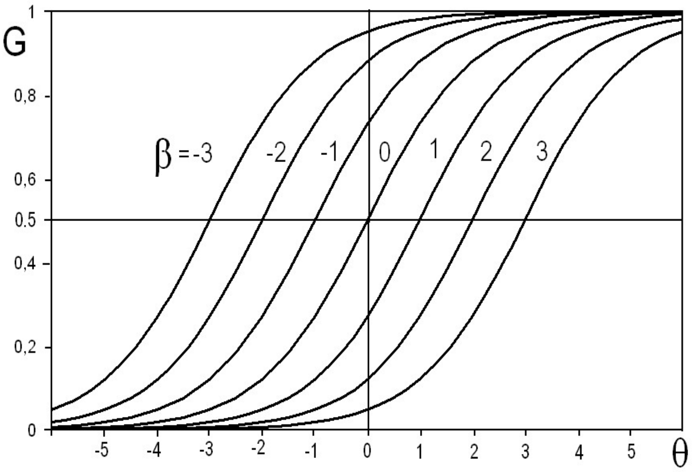

Рис. 1. *Зависимость вероятности правильного ответа испытуемого от сложности задачи и его подготовленности*

Использование модели Раша обеспечивает независимость оценок заданий от испытуемых и оценок испытуемых от параметров заданий. Из приведенных зависимостей видно, что чем выше уровень подготовленности θ испытуемого, тем выше вероятность успеха в том или ином задании. Видно, что при $\theta = \beta$ вероятность правильного ответа G равна *0,5*.

III. Определение уровней подготовленности экспертов и сложности заданий с использованием малой выборки тестовых вопросов. Необходимым условием для измерения уровня подготовленности, не зависящего от сложности задачи, является наличие тестовой базы, состоящей из большого числа вопросов разного уровня сложности. Как показывают оценки, исходя из закона больших чисел и результатов компьютерного моделирования, приведенных ниже, чтобы охватить диапазон измерения компетентностей от –7 до 7 логит с точностью хотя бы ±0.1, необходимо заранее подготовить тестовую базу порядка 1000 вопросов с уровнями сложности от –8 до 8 логит. Здесь возникает чисто методологическая трудность. Как правило, существующие базы тестов не охватывают такие количества вопросов, а, главное, они предназначены только на какую-либо сравнительно однородную группу экспертов. С другой стороны, трудно себе представить, чтобы эксперты смогли выделить время для ответа на сотни вопросов. Если у нас имеется большая совокупность тестовых вопросов с известными значениями β, расположенная по порядку возрастания сложности, то для измерения компетентности эксперта можно применять следующую процедуру, снижающую число задаваемых вопросов K от 50 до 100.

Допустим, у нас имеется база из 1000 пронумерованных тестовых вопросов. Будем задавать эксперту последовательность вопросов, начиная с некоторого номера *P* с шагом *L* (величина *L* может быть в диапазоне от 10 до 20, а *P* – случайное число от 1 до [*L/2*]), запоминая при этом номер вопроса *I*, в котором была допущена первая ошибка. Задаем вопросы до тех пор, пока эксперт не допустит подряд *N* ошибок (*N=3÷5*). Допустим, что номер первого вопроса, в котором эксперт ошибся – *J*. Определяем середину этого диапазона $M = [\frac{I+J}{2}]$ и границы последовательности номеров задаваемых вопросов от $i_1 = M - [\frac{K}{2}]$ до $i_2 = M + [\frac{K}{2}]$ с шагом 1.

В качестве значения компетентности эксперта θ рассчитывается средне-взвешенное число

$$\theta = \frac{1}{k}\sum_{i=i_1}^{i_2} \delta_i \beta_i , \quad (2)$$

где $\delta_i = 0$, если ответ *i* неправильный или $\delta_i = 1$, если правильный, *k* –число правильных ответов.

Аналогичным образом определяется сложность нового, не имеющегося в базе, задания. Допустим, что у нас в базе данных имеется список экспертов, с известными значениями θ_j, расположенных в порядке возрастания. По процедуре, изложенной выше, определяется последовательность экспертов от j_1 до j_2 с шагом 1, решающих правильно новое задание с вероятностью примерно равной 0.5. Сложность нового задания можно определить при этом следующим образом:

$$\beta = \frac{1}{k}\sum_{j=j_1}^{j_2} \delta_j \theta_j . \quad (3)$$

IV. Применение краудсорсинга для составления базы тестовых вопросов. Как известно, составление тестовой базы с большим количеством вопросов является весьма трудоемкой и затратной процедурой. Здесь может помочь использование краудсорсинга, проводимого среди экспертов в той области знаний, в которой предполагается в дальнейшем использование сетевого интеллекта протестированных экспертов.

Эксперты разбиваются на группы, допустим, по семь человек, и каждая группа, работая по технологии метода эволюционного согласования решений [6], составляет тестовые вопросы с ответами. Предполагается, что каждый участник генерирует более двух таких вопросов с ответами, и они проверяются, дополняются или отвергаются коллективным разумом группы. В конце итерационного процесса остается, допустим, десять лучших тестовых вопросов. Одновременно группа экспертов определяет примерную сложность вопросов в логитах. Если в процессе таким образом организованного краудсорсинга было организовано, допустим, 80 групп, то на выходе мы можем получить 800 тестовых вопросов разного уровня сложности с правильными ответами.

Поскольку сложности вопросов были оценены приближенно, то пользоваться этими оценками для измерения компетентностей экспертов нельзя. Эти оценки могут быть использованы только при проведении описанной выше адаптационной процедуры тестирования экспертов для того, чтобы снизить количество предлагаемых вопросов каждому эксперту. При этом необходимо в процессе накопления статистики правильных ответов переупорядочивать список вопросов по сложности, делая ранжирование вопросов по сложности более точным.

V. Сертификация экспертов и тестов с использованием краудсорсинга. Имея обширную базу тестовых вопросов различной сложности можно построить следующую процедуру определения абсолютных компетенций экспертов и получения уточненных значений сложностей тестовых вопросов.

Предположим, что у нас есть подготовленная заранее база тестовых вопросов из 800 вопросов разной степени сложности, предварительно проранжированных по степени сложности неким жюри, компетентность которого значительно выше компетентности тестируемых экспертов, и коллектив из 500 экспертов, подлежащих сертификации,

По адаптационной методике, описанной выше, каждому из экспертов предлагается по 50 вопросов из тестовой базы, причем эксперту задаются вопросы, в формулировании которых он не принимал участия и, следовательно, априори не знает правильных ответов. Фиксируются его правильные и неправильные ответы. В процессе накопления информации обо всей дополняющейся совокупности правильных и неправильных ответов всего коллектива экспертов происходит уточнение порядка распределения вопросов по сложности. Для этого периодически происходит сортировка списка вопросов по количеству правильных ответов. После окончания процесса тестирования, когда каждый из пятисот экспертов ответил на свои 50 вопросов, мы получаем двумерную таблицу $T_{j,i}$, где j – номер эксперта, а i – номер тестового вопроса. Если j-й эксперт ответил на i-й вопрос правильно, то в соответствующее поле таблицы $T_{j,i}$ записываем единицу, если неправильно, то нуль. Если данный вопрос эксперту не предлагался, то в это поле ставим прочерк.

Рассмотрим методику, которая при обработке полученной таблицы позволяет однозначно определить уровни подготовленности экспертов и сложности тестовых вопросов. Отметим, что мы уже имеем первое приближение для уровней сложности тестовых вопросов, они были определены коллективным интеллектом экспертов на стадии формирования базы тестов. Первое приближение может быть также найдено группой экспертов, обладающих высокой компетентностью.

Следуя Рашу[5], в каждом i-м столбце подсчитываем сумму правильных ответов всех экспертов S_i и записываем в нижней части таблицы, а в каждой j-й строке подсчитываем сумму правильных ответов Q_j и записываем справа:

Таблица 1

$T_{j,i}$	1	2	3	4	…	m	
1	0	1	-	0	…	-	Q_1
2	1	-	0	1	…	0	Q_2
3	0	-	0	0	…	0	Q_3
4	-	0	1	-	…	1	Q_4
…	…	…	…	…	…	…	…
n	1	0	0	1	…	0	Q_n
	S_1	S_2	S_3	S_4	…	S_m	

Итак, в таблице 1 находится информация о результатах тестирования n экспертов на m тестовых вопросах.

Далее делаем двойную сортировку таблицы $T_{j,i}$ по строкам в соответствии с полученными значениями Q_j и столбцам – в соответствии со значениями S_i таким образом, чтобы слева оказались самые сложные тестовые вопросы, а сверху – эксперты со слабой подготовленностью. В результате образуется таблица 2, в которой в верхней левой части преимущественно будут расположены нули, а в правом нижнем углу – единицы.

Таблица 2

$T_{j,i}$	β_1	β_2	β_3	β_4	…	β_m
θ_1	0	0	-	0	…	-
θ_2	0	0	0	-	…	0
θ_3	0	-	0	0	…	1
θ_4	0	0	-	1	…	1
…	…	…	…	…	…	…
θ_n	-	0	1	1	…	1

После проведения двойной сортировки перенумеруем экспертов и тестовые вопросы в соответствии с полученным порядком, а в верхнюю строчку таблицы 2 поместим значения первого приближения для сложностей. Значения уровней подготовленностей экспертов пока не определены.

Исходя из данных таблицы 2 для большинства клеток, расположенных на пересечении *j*-х строк и *i*-х столбцов можно рассчитать величину вероятности правильного ответа *j*-го эксперта на *i*-й вопрос как отношение числа правильных ответов в некоторой окрестности данной клетки к числу клеток этой окрестности:

$$G_{j,i} = \frac{1}{(2l+1)^2} \sum_{p=j-l}^{j+l} \sum_{r=i-l}^{i+l} T_{p,r} \delta_{p,r} , \qquad (4)$$

здесь l – размер окрестности, $\delta_{p,r}$ *=1*, если ответ правильный и $\delta_{p,r}$ *=0* в остальных случаях. Здесь следует отметить, что по формуле (4) значения вероятностей правильных ответов можно подсчитать только для областей от $j= l+1$ до $n–l$ и $i=l+1$ до $m–l$. В областях, лежащих за пределами этой части прямоугольной таблицы, значения $G_{j,i}$ можно определить с меньшей точностью, уменьшая размер окрестности. Причем очевидно, что значения $G_{j,i}$ в левой верхней части таблицы, как правило, равны нулю, а в правой нижней части – единице. В правой верхней части таблицы и левой нижней части таблиц можно ставить прочерки. Итак, мы получили окончательную таблицу с вероятностями правильных ответов, с неизвестными значениями уровней подготовленности экспертов и первыми приближениями для уровней сложности вопросов в следующем виде:

Таблица 3

$G_{j,i}$	β_1	β_2	β_3	β_4	…	β_{m-2}	β_{m-1}	β_m
θ_1	0	0	0	$G_{1,3}$	…	-	-	-
θ_2	0	0	$G_{2,3}$	-	…	$G_{2,m-2}$	-	-
θ_3	0	-	$G_{3,3}$	$G_{3,4}$	…	$G_{3,m-2}$	$G_{2,m-1}$	-

θ_4	$G_{4,1}$	$G_{4,2}$	$G_{4,3}$	$G_{4,4}$	…	$G_{4,m-2}$	$G_{4,m-1}$	$G_{4,m}$
…	…	…	…	…	…	…	…	…
θ_{n-2}	-	$G_{n-2,2}$	$G_{n-2,3}$	$G_{n-2,4}$	…	$G_{n-2,m-2}$	$G_{n-2,m-1}$	1
θ_{n-1}	-	-	$G_{n-1,3}$	$G_{n-1,4}$	…	$G_{n-1,m-2}$	1	1
θ_n	-	-	-	$G_{n,4}$	…	1	1	1

Величину $G_{j,i}$ при *α=1* в соответствии с (1) можно выразить формулой

$$G_{j,i} = \frac{1}{1+e^{\beta_i-\theta_j}} . \tag{5}$$

Логарифмируя, получим расчетное выражение

$$\beta_i - \theta_j = \ln\frac{1-G_{j,i}}{G_{j,i}} = C_{j,i} . \tag{6}$$

Исходя из (6) рассчитаем таблицу значений $C_{j,i}$, оставляя имеющиеся в таблице 3 прочерки и заменяя нули и единицы также на прочерки, поскольку делить на нуль нельзя, а логарифм нуля не существует.

Далее строим итеративную процедуру нахождения величин θ_j и β_i, учитывая, что первое приближение для β_i известно:

1. Находим приближение для всех θ_j, выбирая из таблицы $C_{j,i}$ все заполненные значения по индексу *i*
   $$\theta_j = \frac{1}{m_j}\sum_i \beta_i - C_{j,i} ,$$
   здесь m_j - количество заполненных значений в строке *j*.
2. Находим среднее значение $\theta_s = \frac{1}{n}\sum_{j=1}^{n}\theta_j$.
3. Вычитаем его из всех θ_j:
   $$\theta_j = \theta_j - \theta_s .$$
4. Вычисляем следующее приближение для уровней сложности: $\beta_i = \frac{1}{n_i}\sum_j \theta_j + C_{j,i}$,
   здесь n_i - количество заполненных значений в столбце *i*.
5. Переходим на пункт 1 до тех пор, пока итеративный процесс вычисления θ_j и β_i не сойдется.

Полученная таким образом база тестовых вопросов может быть использована для сертификации новых экспертов. Процедура адаптивного измерения подготовленности экспертов была изложена выше. Аналогичным образом группа сертифицированных экспертов может определять уровни сложности новых тестовых вопросов.

VI. Испытания технологии сертификации экспертов и тестов с помощью компьютерной модели. Для проверки предложенной методики сертификации экспертов и определения уровней сложности тестовых вопросов были проведены эксперименты на компьютерной модели. Компьютерное моделирование осуществлялось следующим образом.

С использованием генератора случайных чисел генерируются таблицы подготовленностей экспертов θ_j и уровней сложностей тестовых вопросов β_i, причем эти распределения имеют существенную нелинейность. Одна из реализаций этих распределений приведена на рис. 2.

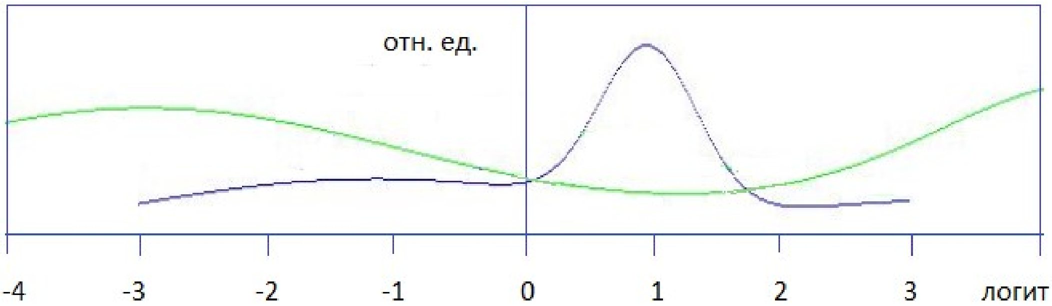

Рис.2. *Относительное распределение экспертов по уровню подготовленности и тестовых вопросов по сложности*

Уровни сложности в рассматриваемом примере меняются от -4 до 4 логит, а уровни подготовленности от -3 до 3 логит.

Исходя из заданных распределений случайным образом были сгенерированы таблицы значений подготовленности экспертов θ_j для *j=1,2,3,…500* и уровней сложности заданий β_i для *i=1,2,3,…,800.* Виртуальные эксперты были подвергнуты «тестированию» - они заполняли таблицу $T_{j,i}$ следующим образом. В ячейку таблицы с координатами *j,i* записывалась единица, если случайная величина, генерируемая компьютером в диапазоне от 0 до 1, оказывалась меньше, чем рассчитываемая по формуле Раша (4), иначе записывался нуль.

Далее вычисления проводились по методике, описанной выше. В качестве первого приближения для уровней сложности β_i выбирались сгенерированные случайным образом арифметические прогрессии.

Результаты восстановления исходных таблиц представлены на рис.3.

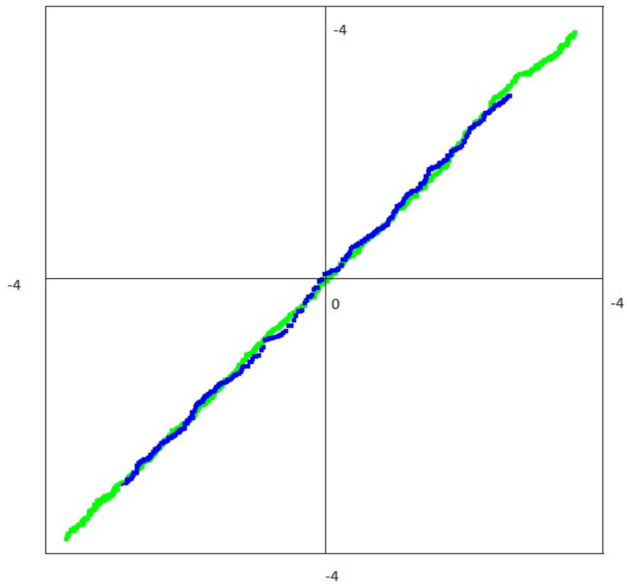

Рис. 3. *Результаты расчетов одного из вариантов*

На этом рисунке по оси абсцисс отложены значения исходных таблиц, а по оси ординат – восстановленные значения, и те и другие – в логитах. Светлая кривая соответствует значениям β_i, темная – θ_j. Погрешности восстановления в абсолютных единицах не превышают 0.12 логит. Результаты расчетов, как показали компьютерные восстановления исходных зависимостей, не зависят от первого приближения. Таким образом, можно сказать, что в рамках модели Раша становится возможным введение абсолютной шкалы измерения подготовленностей специалистов и меры сложностей задач.

VII. Краткое описание метода эволюционного согласования решений. В [9] приведено описание новой информационной технологии коллективного решения интеллектуальных задач с использованием метода эволюционного согласования решений(МЭС). Метод представляет собой модифицированную для компьютерных сетей технологию Дельфи, в которой координаторами групповой работы экспертов выступают генетические алгоритмы.

Определим этот метод следующим образом. МЭС [10] – способ организации коллективной работы экспертов над проектом с заранее заданной целью по правилам, основанным на принципах классического генетического алгоритма. Проект разбивается на отдельные слоты, подлежащие заполнению. Эксперты в соответствии со своими знаниями (умениями) заполняют слоты правильными или неправильными ответами, либо, если они не знают ответ, оставляют их незаполненными. Правила по организации работы экспертов и их взаимодействия выглядят следующим образом:

1) сформулированы цели проекта;

2) определяется состав экспертов и способ их взаимодействия;

3) задаётся каркас проекта – перечень слотов, подлежащих заполнению;

4) находятся первые варианты решений, возможно неполные;

5) проводится обмен вариантами решений;

6) проверяются критерии окончания работы – слот считается заполненным, если более половины экспертов заполнили его одинаковым образом;

7) из полученных решений составляются новые решения (скрещивание);

8) в новые решения вносятся изменения (мутация);

9) осуществляется переход на п.5.

В соответствии с правилами взаимодействия разрабатываются инструкции для коллективной работы с учётом особенностей конкретной задачи, коммуникационной среды, способностей и квалификации интеллектуальных агентов.

VIII. Определение относительной стоимости решения задачи в зависимости от ее сложности. Как будет показано в этом разделе, совместное применение модели Раша и МЭС позволяет принципиально решить проблему квалиметрии сложности задач и вклада экспертов в коллективный проект. Как видно из описания МЭС, креативные способности экспертов, используемые ими при коллективной работе над проектом, состоящим из слотов одинаковой сложности, можно свести к четырем параметрам:

G_r - вероятность правильного заполнения слота проекта экспертом на начальном этапе (этап генерации идей),

G_n - вероятность ошибки на этом этапе,

E_r - вероятность правильной экспертизы предъявляемых на проверку чужих вариантов слотов проекта на этапах согласования решений,

E_n - вероятность ошибки на этих этапах.

Эти параметры могут быть измерены при тестировании экспертов на специально сконструированных тестах с вопросами различной степени сложности таким же образом, как это было представлено в разделах V и VI.

Способности эксперта к генерации идей проверяются на тестах с открытыми вопросами – эксперт должен вписать свой ответ в пустую графу теста и, если ответ совпадает с ключом, то засчитывается правильный ответ, если же не совпадает, то засчитывается неправильный. Если эксперт не знает ответа на поставленный вопрос, то он оставляет соответствующую графу незаполненной. Способности эксперта к экспертизе чужих решений проверяются на специально сконструированных вопросах закрытого типа, когда эксперту предлагаются варианты ответов, среди которых могут быть правильные. Эксперт должен дать ответ - есть ли среди предложенных вариантов правильный ответ, и, если он есть, то указать его. Ответ эксперта, верно определившего отсутствие правильного ответа, также считается правильным. Эксперт может дать ответ «не знаю». Здесь уместно будет отметить, что для большего успеха коллективной работы, экспертам выгоднее в случае сомнений давать ответ «не знаю», чем отгадывать правильный ответ. Это связано с тем, что каждый неправильно заполненный слот уменьшает вероятность правильного ответа группы и снижает рейтинг эксперта.

Зависимость вероятности правильного ответа эксперта на этапе генерации идей G_r от степени подготовленности эксперта θ_G и сложности вопроса β следуя Рашу можно записать в виде

$$G_r = \frac{1}{1+e^{\beta-\theta_G}}. \qquad (7)$$

Степень подготовленности эксперта θ_G определяется, как это было сказано выше, на тестах с открытыми вопросами.

Вероятность неправильного ответа эксперта G_n на этом этапе в зависимости от сложности задачи и подготовленности эксперта можно определить также из проверки на тестах с открытыми вопросами. После статистической обработки результатов тестирования было установлено, что эта зависимость может быть описана следующим образом:

$$G_n = A_G \exp(-(\beta-\theta_G)^2). \qquad (8)$$

Здесь A_G - значение вероятности неправильного ответа на вопрос при $\theta_G = \beta$.

Действительно, анализ формулы (8) показывает, что при ответе на простые вопросы, когда $\theta_G >> \beta$, вероятность неправильного ответа близка к нулю. С другой стороны, в случае решения сложных задач, когда $\theta_G << \beta$, вероятность неправильного ответа также должна стремиться к нулю, поскольку эксперт с вероятностью, близкой к единице будет давать ответ «не знаю».

Из анализа результатов тестирования большой группы экспертов с использованием вопросов закрытого типа в рамках модели Раша, было получено, что зависимость вероятности правильной экспертизы E_r на этапах согласования в зависимости от сложности задачи β и подготовленности эксперта к экспертизе проектов θ_E можно описать выражением:

$$E_r = \frac{1}{1+e^{\beta-\theta_E}}. \qquad (9)$$

Аналогично (8) зависимость вероятности неправильной экспертизы E_n на этапах согласования в зависимости от сложности задачи β и подготовленности эксперта к экспертизе проектов θ_E можно представить в виде:

$$E_n = A_E \exp(-(\beta - \theta_E)^2). \qquad (10)$$

Исходя из представленных выражений и возможности расчета компьютерным моделированием вероятности правильного ответа, полученного группой экспертов в зависимости от сложности задачи и их креативных характеристик, можно построить методику квалиметрического обеспечения системы определения относительной стоимости выполнения задач и справедливой системы оплаты труда специалистов, решающих интеллектуальные задачи в составе группы или индивидуально.

Для построения этой методики введем понятие идеального эксперта. Идеальным экспертом будем считать такого эксперта, у которого зависимость вероятности правильного решения задачи от сложности определяется выражением (7), а остальные характеристики таковы: $G_n = 0$, $E_r = 1$ и $E_n = 0$. Будем называть гарантированным решением какой-либо задачи ее правильное решение с вероятностью не ниже 0.999. Например, из (7) можно получить, что для гарантированного решения задачи сложностью β логит требуется работа одного специалиста с подготовленностью θ_r не ниже $\beta+7$ логит.

С использованим компьютерной модели МЭС можно рассчитать какое количество идеальных экспертов квалификации θ_r нужно для гарантированного решения задач разной сложности. Результаты расчетов приведены в таблице 4.

Таблица 4.

n	*1*	*2*	*3*	*4*
Сложность задачи β	$\theta_r - 1$	θ_r	$\theta_r + 1$	$\theta_r + 2$
Количество экспертов M_n	5	10	22	55
Отношение M_n / M_{n-1}		2	2.2	2.5

Самое меньшее отношение числа экспертов одинаковой квалификации M_n, гарантированно решающих задачу, к числу таких же экспертов, решающих более простую задачу сложностью на один логит меньше, расположено во втором столбце. Следовательно, если стоимость решения задачи сложностью в нуль логит принять за единицу, то стоимость решения более сложной задачи в один логит будет в два раза больше. Естественно, фирма, нанимающая специалистов для решения задач, выберет именно это соотношение. Следовательно, относительная стоимость *C* решения задачи сложности β должна составить величину

$$C = C_0 2^{\beta}, \qquad (11)$$

где C_0 - цена, установленная за решение задачи сложностью в нуль логит.

Соответственно, специалисту с квалификацией $\theta_G = \beta$, $G_n = 0$, $E_r = 1$ и $E_n = 0$, принимавшему участие в гарантированном решении задачи сложностью β в составе коллектива из 10 человек, нужно заплатить

$$Z = 0.1 C_0 2^{\beta}. \qquad (12)$$

Для того, чтобы оценить уровень притязаний реального эксперта, обладающего измеренными в результате тестирования параметрами θ_G, A_G,

θ_E и A_E, необходимо с помощью компьютерной модели рассчитать, сколько нужно экспертов такой же квалификации для гарантированного решения задачи сложностью $\beta = \theta_G$. Далее с использованием компьютерной модели нужно рассчитать величину $\theta_G^{ид}$ для идеального эксперта исходя из ранее полученного количества реальных экспертов и по формуле (12) для $\beta = \theta_G^{ид}$ рассчитать, на какую сумму оплаты своего труда может претендовать специалист. В таблице 5 приведены результаты расчетов величины $\theta_G^{ид}$ для специалистов разных квалификаций. Из анализа этой таблицы видно, что чем больше величины вероятностей неправильных решений, тем больше снижается величина $\theta_G^{ид}$ по сравнению с θ_G. Величина $\theta_G^{ид}$ является по сути интегральной характеристикой креативных способностей специалиста.

Таблица 5.

№	θ_G	A_G	θ_E	A_E	β	M(β)/Z	M(β+1)/Z	M(β+2)/Z	$\theta_G^{ид}$
1	-1,6	0,08	-0.4	0,1	-3	4/31,2	9/27,8	17/29,4	-1,68
2	-0,4	0,11	0,7	0,12	-2	4/62,5	9/55,6	17/58,8	-0.68
3	0,5	0,31	1,5	0,24	-1	5/100	32/31,2	89/22,5	0,00
4	1,3	0,22	2,3	0,31	0	6/167	25/80	62/64,5	0,74
5	2,2	0,16	3,5	0,15	1	5/400	15/267	23/348	2,00
6	2,8	0,24	3,8	0,33	1	4/500	12/333	56/142	2,32
7	3,8	0,35	4,5	0,42	2	4/1000	20/400	309/52	3,32
8	4,3	0,24	5,3	0,11	3	6/1333	27/593	32/1000	3,73
9	5,1	0,12	6,2	0,24	4	6/2667	14/2285	32/2000	4,73
10	6,1	0,22	7,6	0,21	5	7/4571	27/2370	29/4413	5,51

В этой таблице также приведены результаты расчетов «заработка» Z (при $C_0 = 1000$ у.е.) десяти разных специалистов при решении задач разной сложности и сколько таких специалистов нужно для их гарантированного решения. Видно, что специалисты, объединяясь в группы, могут выбирать себе задачи определенной сложности, чтобы максимизировать свой заработок.

IX. Заключение. В результате проделанной работы можно сделать вывод, что использование МЭС, модели Раша, а также итерационного метода определения уровней подготовленности экспертов и сложности тестовых вопросов, позволяет решить проблему квалиметрии в тестировании и решении интеллектуальных задач. Становится возможным однозначное и объективное измерение креативных способностей специалистов, нахождение интегральной оценки качества работы специалиста, относительных величин стоимости решения задачи и размера справедливой оплаты труда специалиста. Поскольку технология достаточно проста и малозатратна, то сообщество экспертов в определенной области человеческой деятельности («цех»), используя краудсорсинг и предлагаемую технологию, может самостоятельно провести разработку тестовых материалов и провести самосертификацию своего сетевого сообщества. В дальнейшем сертифицированные таким образом эксперты могут участвовать в разного рода индивидуальных и коллективных проектах с прогнозируемым результатом.

Одним из важнейших результатов компьютерного моделирования является формулирование следующих утверждений:

1. При увеличении сложности задачи на один логит цена ее гарантированного решения возрастает по меньшей мере вдвое.
2. Гарантированное решение задачи имеет минимальную цену, если подготовленность экспертов группы и уровень сложности задачи отличаются не более чем на один логит.

Благодарности. Авторы признательны фонду РФФИ, профинансировавшему данное исследование в рамках проектов 13-07-00958 «Разработка теории и экспериментальные исследования новой информационной технологии самоуправляемого краудсорсинга» и 13-07-00272 «Методика автоматического формирования ассоциативных портретов предметных областей на основе естественно-языковых текстов больших объемов для систем извлечения знаний».

ЛИТЕРАТУРА

1. Jeff Howe. The Rise of Crowdsourcing. Wired. 2006, p.1-4
2. http://sberbank21.ru/crowdsourcing.html
3. Condorcet, marquis de (Marie-Jean-Antoine-Nicolas de Caritat) (1785), Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix. Imprimerie Royale, Paris.
4. В. Н. Дружинин. Экспериментальная психология: Учебник для вузов / В. Н.. – 2-е изд., доп. – СПб.: Питер, 2003. – 319 с
5. Rasch G. Probabilistic Models for Some Intelligence and Attainment Tests /Expanded Edition, with Foreword and Afterword by B.D. Wright. Chicago: University of Chicago Press, 1980.
6. Аванесов В.С. Применение тестовых форм в Rasch Measurement // Педагогические Измерения № 4, 2005, С.3-20;
7. Аванесов В.С. Метрическая система Георга Раша // Педагогические Измерения №2, 2010, С. 57-80.
8. Аванесов В.С. Три источника становления метрической системы Георга Раша (RM) // Педагогические Измерения №4, 2011, С. 18-29.
9. В.И. Протасов. Конструирование метасистемных переходов. –изд. «Институт физико-технической информатики», 2009 г. 197 с.
10. Протасов В.И. Применение сетевого метода эволюционного согласования решений в управлении проектами. Управление проектами и программами. –М., изд. Grebennikov, 2011, т. 1(25). с. 22-35.